

ID AF017250 standard; RNA; VRT; 5511 BP.
 XX
 AC AF017250;
 XX
 SV AF017250.1
 XX
 DT 06-JAN-1999 (Rel. 58, Created)
 DT 03-MAR-2000 (Rel. 62, Last updated, Version 2)
 XX
 DE *Oreochromis aureus* vitellogenin precursor (Vtg1) mRNA, complete cds.
 XX
 KW .
 XX
 OS *Oreochromis aureus*
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
 OC Cichlidae; Oreochromis.
 XX
 RN [1]
 RP 1-5511
 RA Lim E.H., Lam T.J., Ding J.L.;
 RT "Cloning of Full-Length *Oreochromis aureus* Vitellogenin cDNA and its
 RT Deduced Primary Structure";
 RL Unpublished.
 XX
 RN [2]
 RP 1-5511
 RA Lim E.H., Lam T.J., Ding J.L.;
 RT ;
 RL Submitted (05-AUG-1997) to the EMBL/GenBank/DDBJ databases.
 RL School of Biological Sciences, National University of Singapore, 10 Kent
 RL Ridge Crescent 119260, Singapore
 XX
 DR SPTREMBL; Q9YGK0; Q9YGK0.
 XX
 FH Key Location/Qualifiers
 FH
 FT source 1..5511
 FT /db_xref="taxon:47969"
 FT /sex="female"
 FT /organism="*Oreochromis aureus*"
 FT /tissue_type="liver"
 FT /dev_stage="estradiol-induced adult"
 FT /clone="pOaVtg1"
 FT 5'UTR 1..13
 FT /gene="Vtg1"
 FT sig_peptide 14..61
 FT /gene="Vtg1"
 FT CDS 14..5380
 FT /codon_start=1
 FT /db_xref="SPTREMBL:Q9YGK0"
 FT /gene="Vtg1"
 FT /function="egg yolk protein precursor"
 FT /product="vitellogenin precursor"
 FT /protein_id="AAD01615.1"
 FT /translation="MRVLVLALAVALAVGDQSNLAPGFASVKTYMYKYEAVLMGGGLPEE
 GLARAGVKIRGKVLIISATSANDYILKLVDPQLLEYSGIWPKDPFHPATKLTTALATQLS
 TPIKFEYTNGVVGRILAAPPGVSTTVLNIYRGIIINLLQLNVKTQNVYEMQESGAHGVCK
 FT TNYVIREDARAERIHLTKTSDLNHCQEIKMKAIGLEHVEKCHDCEARGKSLKGTASYN"

FT IMKPAPSGSLIMEAVAREVIEFSPFNILNGAAQMESKQILTFLDIENTPVDHARYTYVH
 FT RGSLQYEHGSEILOQTPIHLLRVTHAEAQIVSTLNHLVASNVAKVHEDAPLKVELIQVM
 FT RVARFETIESLWAQFKSRPDHYWLLNAVPHIRTHAALKFLEKLLANELSETEAAMAL
 FT LECLHSVTADQKTIELVRSLAENHRVKRNRNAVNEIVMLGWGTVISRFCKAQPSCSSDLV
 FT TPVHRQVAEAVETGIDQLTVTLKCLDNAGHPASIKTIMKFLPGFGSAAARVPLKVQVD
 FT AVLALRIIAKREPKMVQEIAAQLLMEKHLHAELRMVAAMVLFETKLPVGLAASISTALI
 FT KEKNLQVVSFVSYSYMKAAMKTSPDHVSVAACNVALRFLNPKLGRNFRYSRAFHVDT
 FT YNNAWMMGAAASAVLINDAATVLPROMIMAKARTYMAGAYVDAFEVGVREGIQEALLKR
 FT RHENSENADRITKIKQAMRALEWRANPSSQALASMYVKVFGQEIAFANIDKSKVDQLI
 FT QFASGPLRNVRDADVNSVLSGYATHFAKPMILLGEELRLILPTTVGLPMEISLITSAVTAA
 FT SVDVQATVSPPLPVNRYVSQLLESDIQLRATVAPS LAMQTYAFMGVNTALIQAAMVTKA
 FT KVYTAVPAQIKARIDIVKGNLKVEFLSLQGINTIASAHAETVIAARNVEDLPAARSTPL
 FT ISSETASQLSKASLNSKISRMASSVTGMSASSEIIIPADLPSKIGRKMKLPKTYRKKIR
 FT ASSRMLGFKAYAEIESHNAAYIRDCPLYALIGKHAASVRIAPASGPVIEKIEVEIQVGD
 FT KAAENMIKAIDMSEEEAELEDKNVLLKIKKILAPGLKNTTSSSSSSSSSSSSSNKS
 FT SSSSSRSSSSQSSSSRSRHSRSRSRKSQSSSSQSSRSPSSSSSSSSSSRSSSSRSS
 FT SRSSSSRSSRSRTKMADIVAPIITTSTRVSSSSRSASNSSSSASYLLSSSKRRSR
 FT SSSSSSSSSSSSSSSSSSSSSKNSKRSKSSNSKSSSSRSSRSAQSKQQLLALKFRK
 FT NHVHRHAISTQRGSSHSSARSFDISYNKAKYLANTLTPAMSIAIRAVRDHKVQGYQLA
 FT AYLDKQTNRQLIFARVAEKDNWRICADIVQLSSHKLMAKTAWGAECKQYSTMIVAETG
 FT LLGHEPAARLKLTWDKLPGSIKYAKRALKSIVPIAQEGYGVNYAKAKNPRNQIKLTVAV
 FT ATETSMNIVLNTPKAIVYKRGVCLPVALPIGNTAAELQATRDNWADKMSYLVTKANAVE
 FT CSLINNTLTTFNNRKARDELPHSCYQVLAQDCTPELKFMVLLKKDQIQDQNQINVKISD
 FT IDVDMYRKNNAIAVMVNGVEIPNSNLPLYHPSGNIHIRQSNEGITALNAPSHGLQEVFLG
 FT FNELRVKADWMKGKTCGACTASGNVGEYRTPSEQVTKDAISYAHSWVLSSNTCRDP
 FT SECSIKQESVKLEKRVIFEGVESKCYSVEPVLQCLPGCIPVRTTVNVGFHCLPSDTTV
 FT DRSGLSSFFEKSIDLRTDAE AHLACRCPQCA"
 FT mat_peptide 62..5377
 FT /gene="Vtg1"
 FT /product="vitellogenin"
 FT misc_feature 3269..3946
 FT /note="encodes putative phosvitin domain"
 FT /gene="Vtg1"
 FT 3'UTR 5381..5511
 FT /gene="Vtg1"
 FT polyA_signal 5465..5470
 FT /gene="Vtg1"
 XX
 SQ Sequence 5511 BP; 1474 A; 1443 C; 1299 G; 1295 T; 0 other;
 acatccacca gccatgaggg tgcttgtact agctcttgct gtggcttcg cagtggggga 60
 ccagtccaaac ttggccccag gattcgctc tggtaagacc tacatgtaca aatatgaagc 120
 ggttcttatg ggcggcctgc ctgaagaggg cctggctcga gctggggta aaatccgggg 180
 caaagttttg atcaagtccaa caagtccaa cgactacatt ctgaagctt tagaccctca 240
 gttgctggag tacagtggca tctggccaa agatccttc catccagcca ccaagctcac 300
 cacagccctg gctactcagc tctcgacacc gatcaagttt gagtatacaa acggcgttgt 360
 cgggagactg gctgcaccc tcggggctc cacaacagtg ctgaatatct acaggggaaat 420
 catcaacctc ctgcagctga atgtaaagaa gacacagaac gtctacgaga tgcacagatc 480
 tggagctcat ggtgtgtca agaccaacta tgtgatcagg gaggacgcga gagccgaacg 540
 cattcatctg accaagacca aggacctgaa ccactgccag gagaaaatca tgaaggccat 600
 cggcttggaa cacgttagaa aatgccatga ttgtgaagct agagggaaaga gcctgaaggg 660
 aactgcttcc tataactaca tcatgaagcc agcacccagt ggttctctga ttatggaggc 720
 tgtcgctaga gaggtcatcg aattctcacc tttcaacatt ttgaatggcg ctgctcagat 780
 ggagtctaag caaattctga ccttccttggaa tattcgagaac acccctgtgg atcatgccag 840
 atacacccat gttcaccgcg gatcccgtca gtatgagcat ggcagcggaa ttctccagac 900
 acccatccat cttctgaggg tcacccatgc cgaggctcag attgtcagca ctctgaacca 960
 cctggtagcc tccaacgtgg ccaagggtcca tgaagatgcc cctctgaagt ttgttgagct 1020
 catccaggtg atgcgtgtgg ccagattga gactatttag tccctctggg ctcagttaa 1080
 atctagaccc gatcacaggt actggttact gaatgctgtc ccccacattc gcactcacgc 1140
 tgcgcttaag ttccatttga agaagctct tgctaatgag ttaagtgaga ctgaagctgc 1200
 tatggctctc ttggaaatgtc tgcactctgt gacagctgac cagaaaacca ttgaacttgt 1260

cagaaggctg	gctgagaacc	acagagtcaa	acgttaacgct	gtgctcaacg	agattgtgat	1320
gctgggctgg	ggcactgtaa	tttccaggtt	ctgttacata	cagccatctt	gctcatctga	1380
tcttgtaca	cctgtacata	gacaagtgc	agaggctgtt	gaaactggtg	acatcgatca	1440
gctcaactgtc	acgctcaaat	gcctggataa	cgctggacat	cctgcttagca	ttaagacaaat	1500
catgaagttc	ctgcctggct	ttggcagtgc	tgctgcccga	gtcccactca	aagttcaggt	1560
tgacgctgtt	ctagccctga	ggagaattgc	aaagagggaa	cccaagatgg	tccaggaaat	1620
agctgctca	ttgctcatgg	aaaagcatct	ccatgcagaa	ctgcgtatgg	ttgctgccccat	1680
ggtgctctt	gagactaaac	tccccgtggg	tctagcagct	agcatttcca	cagccttgat	1740
caaagaaaag	aacgtcagg	tcgttagctt	tgtctactct	tacatgaagg	ccatggccaa	1800
gaccacatcc	cctgaccacg	tttctgtgc	tgcagcatgt	aatgttcctt	tgaggttcct	1860
caaccccaa	ttagggcagac	tgaacttccg	ctacagccga	gccttcat	tggtatcaccta	1920
taacaatgcc	tggatgatgg	gtgctccgc	cagtgcgtc	ttaattaacg	acgctgcaac	1980
cgtgttacca	agaatgatta	tggccaaagc	ccgtacttac	atggccggag	cttatgttg	2040
tgctttttag	gttggagtga	ggactgaggg	aatccaggag	getctttga	aaagacgaca	2100
tgaaaattct	gagaatgcag	acaggatcac	caagattaaa	caagccatga	gagcttttc	2160
tgagtggagg	gctaattcctt	cgagccaggc	cctggctctt	atgtatgtga	aggtcttcgg	2220
acaagaaatt	gcatttgcca	acattgacaa	atccaagggtt	gaccagctt	tccagtttgc	2280
cagtgacact	ttgagaaaacg	tattcagaga	tgctgtgaat	tctgtgtgt	ctggttatgc	2340
aacacatttt	gctaaaccaa	tgctgctcg	tgagctccgt	ctcatccctt	ccaccactgt	2400
tgggttgc	atggagatca	gcctcattac	atccgctgt	actgctgcat	ctgttgcgt	2460
ccaagccact	gtgtcaccac	ctctgcctgt	caactaccga	gttcccagc	ttctggagtc	2520
cgatataccaa	ctgagggcta	cagttgc	aatcttgc	atgcagac	atgcattcat	2580
gggtgtgaac	accgccttaa	tccaggctgc	agtgtatgaca	aaagccaaag	tttacacagc	2640
tgttctcg	cagataaaag	caaggatgt	catgtttaag	ggcaacttga	agtttgagtt	2700
cctgtcttc	caggcattta	acacaatgc	atctgcacat	gcccggacgg	ttgcattgc	2760
aagaaatgtg	gaagacctcc	cagccgcaag	aagcacacca	ctgatctcat	ctgaaactgc	2820
atcacaactt	tcaaaggcct	ctctcaactc	aaagatctcc	aggatggcat	cctctgtgac	2880
tggtggcatg	tctgcgtcat	ctgaaatcat	tcctgctgac	ctgccaagta	agattgggag	2940
gaaaatgaaa	ctccctaaaa	cctacaggaa	gaaaatccgt	gcttcaagca	gaatgtctt	3000
attcaaggcc	tacgctgaga	ttgaatctca	caatgcgc	tacatcagag	actgcctct	3060
ctacgctctg	atcggaaagc	atgctgc	tgttaggatt	gctccagctt	ctggaccagt	3120
cattgagaag	attgaagttg	agattcagg	cgagataaaa	gcagcagaaa	atatgattaa	3180
agcgattgac	atgagcgaag	aggaggaagc	tcttgaggat	aagaatgtcc	tcttggaaaat	3240
caagaaaata	ctggcacctg	gtctcaagaa	caccacatca	tcttcctcca	gctcctccag	3300
ctcctcttca	tccagctcta	gctccaacaa	gtcttcttca	tccagttccc	gctccagcag	3360
ctcccagtc	tccagctctc	gttcccata	gtctcgctc	agaaaatccc	agtctagcag	3420
ctctcagtc	agccgctctc	cctcaagtc	ttcctccctt	tccttcctt	catcatccag	3480
atcttcttcc	aggcatctt	ccagatcatc	ttccagatct	tcttcttaggt	cctcctctcg	3540
ctccagaact	aagatggctg	acattgttgc	tccttattatc	acgacgtcca	ccagagttag	3600
cagttcctcc	agtgcgtac	cctctaacag	ctcttccagc	agtgcgtat	acttgcgtac	3660
ctcatctaaa	agaagaagca	gaagcagaag	cagcagcagc	agtgcgtat	gcagcagcag	3720
cagcagcgt	agcagcagca	gcagcagcag	cagtagcaag	aacagcaagc	gcagcaagag	3780
cagcaacagc	aagtcatcaa	gctctagtc	ctctcggcgc	agtgcgtat	ctaagcaaca	3840
actgcttgc	ttgaagttca	gaaagaacca	cgtccacagg	catgcacatct	ccacacagcg	3900
cggcagcgt	cacagcagtg	cccgacgtt	cgattccatc	tacaataagg	ccaaagtacat	3960
cgctaaacaca	ctcactcctg	ccatgtccat	tgcaatccgt	gccgtgagag	tcgaccacaa	4020
ggtccaggga	taccagctg	cagcttac	ggacaaacag	accaatagac	tgcaatgtat	4080
ttttgcaga	gtcgctgaga	aggacaactg	gagaatctgt	gccgcacattt	tgcaatgtat	4140
ttcgcacaag	ctgtggcca	agactgc	gggtgctgaa	tgcaagcaat	actccacat	4200
gattgttagt	gaaactggc	ttttgggtca	tgagcccgca	gcccgttga	agctgacat	4260
ggacaaactg	ccaggaagca	taaagacta	cgcggaaaggg	gcgttggaaat	ccattgtccc	4320
tattgctaa	gaatatggag	taaactacgc	aaaggccaa	aatcctcgta	atcaaatcaa	4380
actgactgt	gctgttgct	ctgagacaag	catgaatatt	gtgctgaaaca	caccaaaaggc	4440
aatcgtttac	aagcgtgggg	tgtgtctacc	tgttgcttta	ccaattggaa	acactgtgc	4500
cgagctgca	gcaacccggg	acaactggc	tgacaagatg	tccttatttg	ttaccaaaagc	4560
taacgcagt	gaatgctccc	tcatcaacaa	cacactgacc	acattcaaca	acaggaaagc	4620
tagagatgag	ctgcccacact	cgtgctacca	ggtcttggct	caggatgtca	caccagaact	4680
caaattcatg	gttctgtca	agaaagacca	aatacaggat	cagaaccaga	tcaatgttaa	4740
gatttcagac	atcgatgtgg	acatgtatcg	gaagaacaac	gccattgcgg	tgatgtttaa	4800
cggagttgaa	atccctaaca	gcaacctgccc	atacctgc	ccatcaggta	acatacatat	4860

aagacagtca aatgaaggca ttactctcaa tgcacccagc catggcttc aggaggtctt	4920
ccttggcttc aacgagctga gggtaaagt tgcagactgg atgaaaggaa agacttgtgg	4980
tgcctgtgga acggcaagcg gaaaatgtcg agacgagttac cgcacaccca gtgaacaggt	5040
gaccaaggat gccatcagct acgcccactc ctgggttctg tcttcaaaca cctgccgtga	5100
tccctccgag tgttccatca agcaggaatc tgtgaagctg gagaagcggg tcatcttgat	5160
aggtgtggag tccaaatgct actctgttga gcccgtgctg cagtgctgc ccggctgttat	5220
cccaagtgaga accactaccg tcaacgttgg cttdcactgc ctgcccagtg acacaactgt	5280
ggaccgttct ggtctgagca gcttcttga gaagagcatc gacctgaggg atactgcaga	5340
agcccacctg gcctgtcgct gcactcctca gtgtgcttaa tctcattgcc ttccgtcaca	5400
tgacattttg tttctgttt taggcagggtg ttgttaatta gctttggctc aacttgtaaa	5460
cttaaataaa ttggaaaagc atgtgaagca aaaaaaaaaa aaaaaaaaaa a	5511

//

ID AF072686 standard; DNA; VRT; 11453 BP.
 XX
 AC AF072686; X93569; X93570; X93571; Z71336;
 XX
 SV AF072686.1
 XX
 DT 11-AUG-1999 (Rel. 60, Created)
 DT 30-NOV-1999 (Rel. 61, Last updated, Version 2)
 XX
 DE *Oreochromis aureus* vitellogenin gene, complete cds.
 XX
 KW
 XX
 OS *Oreochromis aureus*
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
 OC Cichlidae; Oreochromis.
 XX
 RN [1]
 RP 1-11453
 RX MEDLINE; 99145366.
 RA Teo B.Y., Tan N.S., Lim E.H., Lam T.J., Ding J.L.;
 RT "A novel piscine vitellogenin gene: structural and functional analyses of
 estrogen-inducible promoter";
 RL Mol. Cell. Endocrinol. 146(1-2):103-120(1998).
 XX
 RN [2]
 RP 1-11453
 RA Teo B.Y., Lim E.H., Ding J.L.;
 RT ;
 RL Submitted (18-JUN-1998) to the EMBL/GenBank/DDBJ databases.
 RL Department of Biological Sciences, National University of Singapore, 10
 RL Kent Ridge Crescent, Singapore 119260, Singapore
 XX
 DR SPTREMBL; Q9YGK0; Q9YGK0.
 XX
 CC On Nov 29, 1999 this sequence version replaced gi:1771394
 CC gi:1771393 gi:1771392 gi:1403509.
 XX
 FH Key Location/Qualifiers
 FH
 FT source 1..11453
 FT /db_xref="taxon:47969"
 FT /sex="female"
 FT /organism="*Oreochromis aureus*"
 FT /clone="OaVg1"
 FT /tissue_type="blood"
 FT /dev_stage="adult"
 FT promoter 1..1677
 FT protein bind 141..146
 FT /note="GATA binding site"
 FT /bound_moiety="GATA factor"
 FT protein bind 315..320
 FT /note="GATA binding site"
 FT /bound_moiety="GATA factor"
 FT protein bind 326..338
 FT /note="estrogen response element (ERE); imperfect ERE"
 FT /bound_moiety="estrogen receptor protein"
 FT protein bind 353..358

```

FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT          440..445
FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT          1001..1006
FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT          1146..1158
FT          /note="estrogen response element (ERE); imperfect ERE"
FT          /bound_moiety="estrogen receptor protein"
FT          1200..1205
FT          /note="GATA binding site"
FT          /bound_moiety="GATA factor"
FT          1649..1655
FT          TATA signal
FT          5'UTR
FT          mRNA
FT          join(1678..1732,2065..2085,2193..2344,2441..2692,
FT          2793..2954,3044..3192,3280..3430,3519..3676,3755..3920,
FT          4000..4125,4211..4429,4538..4755,4850..4947,5035..5159,
FT          5288..5392,5502..5621,5709..5892,5990..6083,6177..6362,
FT          6462..6612,6694..6910,7017..7744,7834..7878,7957..8025,
FT          8117..8323,8407..8549,8736..8874,8978..9107,9224..9407,
FT          9497..9591,9686..9780,9906..10071,10213..10374,
FT          10471..10690)
FT          /product="vitellogenin"
FT          CDS
FT          join(1693..1732,2065..2085,2193..2344,2441..2692,
FT          2793..2954,3044..3192,3280..3430,3519..3676,3755..3920,
FT          4000..4125,4211..4429,4538..4755,4850..4947,5035..5159,
FT          5288..5392,5502..5621,5709..5892,5990..6083,6177..6362,
FT          6462..6612,6694..6910,7017..7744,7834..7878,7957..8025,
FT          8117..8323,8407..8549,8736..8874,8978..9107,9224..9407,
FT          9497..9591,9686..9780,9906..10071,10213..10374,
FT          10471..10580)
FT          /codon_start=1
FT          /db_xref="SPTREMBL:O9YGK0"
FT          /note="phospholipoglycoprotein; yolk protein precursor"
FT          /product="vitellogenin"
FT          /protein_id="AAD48085.1"
FT          /translation="MRVLVLALALAVAGDQSNLAPGFASVKTYMYKYEAVLMGGLEE
FT          GLARAGVKIRGKVLSATSDNYILKLVDPQLLEYSGIWPKDPFHPATKLTTALATQLS
FT          TPIKFETYNGVVGRLAAPPGVSTTVLNIYRGIIINLLQLNVKTQNQVEMQESGAHGVCK
FT          TNYVIREDARAERIHLTKTDLNHCQEKIMKAIGLEHVEKCHDCEARGKSLKGTASYNY
FT          IMKPAPSGSLIMEAVAREVIEFSPFNILNGAAQMESKQILTFLDIENTPVDHARYTYVH
FT          RGSLQYEHGSEILQTPHIPLLRTVTHAEAQIVSTLNHLVANVAKVHEDAPLKVFELIQVM
FT          RVARFETIESLWAQFKSRPDHYWLLNAVPHIRTHAALKFLIEKLLANELSETEAAMAL
FT          LECLHSVTADQKTIELVRSLAENHRVKRNAVLINEIVMLGWGTVISRFCKAQPSCSSDLV
FT          TPVHRQVAEAVETGDIDQLTVTLKCLDNAGHPASIKTIMKFLPGFGSAAARVPLKVQVD
FT          AVLALRRRIAKREPDKMVQEIAAQQLMEKHLHAELRMVAAMVLFETKLPVGLAASISTALI
FT          KEKNLQVVSFVSYSYMKAAMAKTTS PDHVSVAACNVALRFLNPKLGRNFRYSRAFHVD
FT          YNNAWMMGAAASAVLINDAATVLPROMMAKARTYMAGAYVDAFEVGVRTEGIQEALLKR
FT          RHENSENADRITKIKQAMRALSEWRANPSSQALASMYVKVFGQEIAFANIDKSVDQLI
FT          QFASGPLRNVRDADVNSVLSGYATHFAKPMILGELRLLILPTVGLPMEISLITSAVTAA
FT          SVDVQATVSPPLPVNRYRSQQLLESDIQLRATVAPSLAMQTYAFMGVNTALIQAAMVTKA
FT          KVYTAVPAQIKARIDIVKGNLKVEFLSLQGINTIASAHAETVAIARNVEDLPAARSTPL
FT          ISSETASQLSKASLNSKISRMASSVTGMSASSEIIIPADLPSKIGRKMKLPKTYRKKIR
FT          ASSRMLGFKAYAEIESHNAAYIRDCPLYALIGKHAASVRIAPASGPVIEKIEVEIQVGD
FT          KAAENMIKAIDMSEEEEALEDKNLLKIKKILAPGLKNTTSSSSSSSSSSSSSSSNKS
FT          SSSSSRSSSSQSSSSRSHRSRSRKSQSSSQSSRSPSSSSSSSSSSSSSSRSSR
FT          SRSSSSRSSRTKMADIVAPIITTSTRVSSSSRSASNSSSSASYLSSSKRRRSR
FT          SSSSSSSSSSSSSSSSSSSSSSSKNSKRSKSSNSKSSSRSSRSAQSKQQLLALKFRK

```


cctggctact	cagctctcga	caccgatcaa	gtttgagtat	acaaacggcg	ttgtcgggag	2580
actggctgca	cctcctgggg	tctccacaac	agtgcgtaat	atctacaggg	gaatcatcaa	2640
cctccctgcag	ctgaatgtaa	agaagacaca	gaacgtctac	gagatgcaag	aggtaaaaaa	2700
aaactcgctt	tcaccaggte	cttcatacaa	aatgcttgaat	caactgacaga	aaataacaac	2760
aatctatatt	ttgattgtat	tttgcgtatct	agtctggagc	tcatgggtgt	tgcaagagcca	2820
actatgtat	cagggaggac	gcgagagccg	aacgcattca	tctgaccaag	accaaggacc	2880
tgaaccactg	ccaggagaaa	atcatgaagg	ccatcggtt	ggaacacgta	gagaaatgcc	2940
atgattgtga	agctgtat	gcattttcta	gttgcattcat	aatctcttctt	atagaagtc	3000
tcatgcgtgc	agtttaaaga	tcgtgttgg	tttatcattt	cagagaggaa	agagcctgaa	3060
gggaactgct	tcctataact	acatcatgaa	gccagcaccc	agtgggtctc	tgattatgga	3120
ggctgtcgct	agagaggta	tcgaatttctc	accttcaac	attttgaatg	gctgctgtca	3180
gatggagtct	aagtatgtgg	tttaaactaa	aataatatgt	ttgattacaa	gcttacaagc	3240
attgtgccac	attgattgaa	cgattatctc	tgctcacagg	caaattctga	ctttcctgaa	3300
tatcgagaac	accctgtgg	atcatgccag	atacacctat	gttcacccg	gatccctgca	3360
gtatgagcat	ggcagcgaga	ttctccagac	acccatccat	cttctgaggg	tcacccatgc	3420
cgaggctcag	gtactgcgc	atttccttgc	agaaagtctt	ttctgaattt	accctgttagc	3480
cacaacactc	tacatctgta	ttgtttgttt	cctcacagat	tgtcagcact	ctgaaccacc	3540
ttgttagcctc	caacgtggcc	aaggtccat	aagatgcccc	tctgaagttt	gtttagctca	3600
tccaggtgat	gcgtgtggcc	agatttgaga	ctattgagtc	cctctgggct	cagtttaaat	3660
ctagacctga	tcacaggtaa	gtgcagctt	gattcagcaa	aggtccacga	aactaacatg	3720
acacggataa	tggcattgtg	ctcctcttcc	ccaggtactg	ttactgaat	gctgtcccc	3780
acattcgcac	tcacgctgca	cttaagttcc	tcatttgagaa	gctccttgc	aatgagttaa	3840
gtgagactga	agctgtat	gctctttgg	aatgtctgca	ctctgtgaca	gctgaccaga	3900
aaaccattga	acttgtcaga	gtaagtcttt	tttaagttt	tcctatataat	ttgcagttagc	3960
atgaaatcag	tttatattt	tgtcttcacc	acttttca	gcctggctga	gaaccacaga	4020
gtgaaacgta	acgctgtgct	caacgagatt	gtgatgctgg	gctggggcac	tgtaattttc	4080
aggttctgta	aagcgcagcc	atctgtca	tctgatctt	tgacagttatg	tattatttga	4140
acatgttttta	ttctgctaaa	gttcagttt	gacatcgtg	actttaatgt	ttttttacat	4200
tcttgacacag	cctgtacata	gacaaggtagc	agaggctgtt	gaaactgggt	acatcgatca	4260
gctcaactgtc	acgctcaat	gcctggataa	cgtggacat	cctgctagca	ttaagacaaat	4320
catgaagttc	ctgcctggct	ttggcagtgc	tgctgcccga	gtcccactca	aagttcaggt	4380
tgacgctgtt	ctagccctga	ggagaattgc	aaagagggaa	cccaagatgg	tgagactaca	4440
ttttctgatt	cctaccatca	ataacaataa	atataacttga	gacattttt	ctgtgcagtg	4500
ttcagttggg	gctatactaa	tcggctttc	tcatttagtgc	cagggaaatag	ctgctcagg	4560
gctcatggaa	aagcatctcc	atgcagaact	gogtatggtt	gctgcccc	tgctctttga	4620
gactaaactc	cccggtggcc	tagcagctag	catttccaca	gccttgcata	aagaaaagaa	4680
cctgcaggc	gttagcttgc	tctactctt	catgaaggcc	atggccaaga	ccacatcccc	4740
tgaccacg	tctgtgtaa	aacctaaatg	attattatgt	ttctaagatg	ctttatatact	4800
atgcatgcta	gtctattgaa	tctcaatata	cccccgtt	tattttaaat	gctgcagcat	4860
gtaatgttgc	cttgagggtt	ctcaacccca	aatttaggc	actgaacttc	cgctacagcc	4920
gagccttca	tgtggatacc	tataacagta	agaagatacc	agataaataa	tttgcgttg	4980
cacacccat	atttgtgcat	gtgggtgtat	gatgtttatc	catttgcct	ccagatgcct	5040
ggatgtatgg	tgctggcc	agtgcgtct	taattaacga	cgctgcaacc	gtgttaccaa	5100
gaatgattat	ggccaaagcc	cgtacttaca	tgccggagc	ttatgttgc	gtttttgagg	5160
taagacgtt	aatgttacg	taagtaacta	aaacagacac	tcagcaatac	aattcaatga	5220
cttcgatatg	taaacaatac	attttcagt	atctgcgtc	gagtgcattc	ataaacgtac	5280
atttcaggtt	ggagtggat	ctgagggat	ccaggaggct	cttttggaaa	gacgacatga	5340
aaattctgag	aatgcagaca	ggatcacca	gattaaacaa	gccatgagag	ctgtaaatgt	5400
aactactctg	gcagagtgaa	aatgggattc	tgagggacat	tttctctat	agagctgttt	5460
tttcttttac	ttgtttttaa	cttaatttat	tccataata	gctttctgag	ttggagggtca	5520
atccttcgag	ccaggccctg	gcctctatgt	atgtgaaggt	cttcggacaa	gaaattgc	5580
ttgccaacat	tgacaatcc	aaggttgc	agcttatcca	ggtagtattt	ctcccacagc	5640
ctcatatgtt	ttcttagtta	tgatgcatta	tttgcattt	aaactaataa	gcaactttgt	5700
ctatacagtt	tgccagtgg	ccttggagaa	acgtattcag	agatgcgtg	attctgtgc	5760
tgtctggta	tgcacacat	tttgctaaac	caatgcgtc	cggtgagctc	cgtctcatcc	5820
ttccacccac	tgtgggttg	cccatggaga	tcagcctcat	tacatccgct	tgactgtcg	5880
catctgttgc	cggtaagttt	cattcagaca	gttacaatgt	ttacagaaag	aaaaaaat	5940
atggatattt	atatggaaaa	tatattgttt	tcttgcgttc	ttttgcgtat	ccaagccact	6000
gtgtcaccac	ctctgcctgt	caactaccga	gttcccagc	ttctgagtc	cgatatccaa	6060
ctgagggtca	cagttgc	aaggtacaca	acagttgag	aagttcttc	tctgcatttt	6120

attgaatgaa ttcaagagata actcttaatg aatgaatatt ttctgtctt atgcagtctt	6180
gccccatgcaga cctatgcatt catgggtgtg aacaccgcct taatccaggc tgcaagtatgt	6240
acaaaaggcca aagtttcac agctgttct gcacagataa aagcaaggat tgacattgtt	6300
aaggggcaact tgaagggtga gttcctgtct ctccaggcga ttaacacaat tgcatctgca	6360
cagtatgtac atgccttgcc actaattgcc atgcattca aagattttg gtcctttcg	6420
gcatctgtt ttattgggtg gggttttt ctctcataca gtgcggagac ggttgcatt	6480
gcaagaaaatg tggaaagacct cccagccgca agaagcacac cactgatctc atctgaaact	6540
gcatcacaaac tttcaaaggc ctctctcaac tcaaagatct ccaggatggc atccctgtg	6600
actggtgccca tggtaagttt ggttcaactcg cctcttattt tcaacattca actgtttgaa	6660
ataaacatgg cgatctctgc tgggtttt cagtcgcgt catctgaaat cattcctgct	6720
gacctgccaa gtaagattgg gaggaaaatg aaactcccta aaacctacag gaagaaaatc	6780
cgtgcttcaa gcagaatgct aggattcaag gcctacgctg agattgaatc tcacaatgccc	6840
gcctacatca gagactgccc tctctacgt ctgatcgaa agcatgctc ttctgttagg	6900
attgctccag gcaagtgcaa gtggtaact tgaagcttta acctaacaaa ttgcääaaac	6960
attgtggaaa agtctgttt ttgttttaa ttattattct ttaatttata ttccagcttc	7020
tggaccagtcc attgagaaga ttgaagttga gattcaggc ggagataaag cagcagaaaa	7080
tatgattaaa gcgattgaca tgagcgaaga ggaggaagct cttgaggata agaatgtcct	7140
cttggaaaatc aagaaaatac tggcacctgg tctcaagaac accacatcat ctccctccag	7200
ctccctccagc tcctttcat ccagctctag ctccaacaag tcttcttcat ccagttcccg	7260
ctccagcagc tcccagtcat ccagctctg ttccatagg tctcgctcca gaaagtccca	7320
gtctagcagc tctcagtcaa gccgctctcc ctcaagctct tcttcctt cctccttctc	7380
atcatccaga tcttcttcca ggtcatctt cagatcatct tccagatctt ctcttaggtc	7440
ctcctctcgc tccagaacta agatggctga cattgttgc tcttattatca cgacgtccac	7500
cagagtggc agttccctca gtcgatcagc ctctaacagc tcctccagca gtgcttcata	7560
cttgctcagc tcatactaaaa gaagaagcag aagcagaagc agcagcagca gtacgacag	7620
cagcagcagc agcagcagta gcagcagcag cagcagcagc agtagcaaga acagcaagcg	7680
cagcaagagc agcaacagca agtcatcaag ctcttaggtc tctcggcgca gtgctcagtc	7740
taaggtgaat catctctctc tcgcagcatg tagaagttgtg tggaaattta tagaatgtgg	7800
agcacttatt ttacaattct tttccttattt cagcaacaac tgcttgctt gaagttcaga	7860
aagaaccacg tccacagggt aattattagg catcacatca aagcacatca tgtcttctgt	7920
ttaagcaggc taaatgtgtg gtttccctgt ttgcagcatg ccatctccac acagcgcggc	7980
agcagtcaca gcagtgcggc cagcttcgat tccatctaca ataagggtat cagtcaaatg	8040
aagatattta cgatgttacg cattaaactg cgaataaaaca taattattaa tatcaacctc	8100
tggtttccac ttttaggcca agtacctcgc taacacactc actcctgcca tgcatttgc	8160
aatccgtgcc gtgagagtcg accacaagg ccagggatac cagctagcag ctacacttgg	8220
caaacagacc aataagactgc agttgatttt tgccagagtc gctgagaagg acaactggag	8280
aatctgtgcc gacattgtgc agctgagttc gcacaagctg atggtaaaca tcatgaactt	8340
atgaggaagt aaccaactaa tattaggtat gatgcttagg atttaactca ctgaaattta	8400
ttttaggcca agactgcctg gggtgctgaa tgcaagcaat actccaccat gattgttagct	8460
gaaactggc ttttgggtca tgagccccca gcccgttga agctgacctg ggacaaactg	8520
ccaggaagca taaagcacta cgcaaagagg taaacttacg gacttttcc gattcataca	8580
tataatcaatgt gcttttattt aaaagctttt gcagacattt tgcagatgga agtgcacagg	8640
aatttattttc ttttgacatc ctctcgaaac tgcatttat actgaaaact gacaaaagct	8700
ttataaaaaca ctcagtaacc ttatgtctt tcttagggcgt tgaaatccat tgcatttctt	8760
gctcaagaat atggagtaaa ctacgcaag gccaagaatc ctcgtaatca aatcaaactg	8820
actgttagctg ttgctactga gacaaggatg aatattgtgc tgaacacacc aaaggtacag	8880
ttttgttgc ccttcttctg tatgcactga aatgtttgc attgcacttt gtcactcaga	8940
gtgtttctaa ctactttctg actgtttaat attccaggca atcggttaca agcgtgggg	9000
gtgtctaccc gttgtttac caattggaaa cactgctgcc gagctgcaag cgaccggg	9060
caactggcgt gacaagatgt cctattttgt taccaaagct aacgcaggta aaaatattt	9120
aattaaacaa aaaacaacaa caacaacaaac aaaaacaggc ttcatgttgc tgcattttgt	9180
cttcagagg ctaacatgtg atttccctt cgtccactc cagttgaatg ctccctcatc	9240
aacaacacac tgaccacatt caacaacagg aaagcttagag atgagctgcc acactcgtgc	9300
taccaggctc tggctcagga ttgcacacca gaactcaat tcatgttct gctgaagaaaa	9360
gaccaaatac aggatcagaa ccagatcaat gttaaaggatt cagacatgta agagcagcc	9420
ggattttatgt ggaaaagaaa taacggtcag cctgtgagcc aaaaataact tatgtatttg	9480
aactgcttct caacagcgt gttggacatgt atcggaaagaa caacgcatt gcggtgatgg	9540
ttaacggagt tgaatccct aacagcaacc tgccataacct gcatccatca gttttttgt	9600
tttggtttgt taatttaaca aaaatattga actgccagat tgcattt gaaaccatgc	9660
ttatatatat ttgcacttac ttttaggttaac atacatataa gacagtcaa tgaaggcatt	9720

actctcaatg caccagcca tggcttcag gaggtttcc ttggctcaa cgagctgagg	9780
gtaatacat aaatgcgtat aattcatata tctaatacgt atattcatta ttttctgagg	9840
caatttccta tttgtatacc tggttaatc tacattgtac taatatctga accttttat	9900
ttcaggtta agttgcagac tggatgaaag gaaagacttg tggtgcctgt ggaacggcaa	9960
gcggaaatgt cgaggacgag taccgcacac ccagtgaaca ggtgaccaag gatgccatca	10020
gctacgccc ctcctgggtt ctgtctcaa acacctgccc tgatccctcc ggtaaagcctt	10080
tgcaatattt ccttgacctt tctctaaaag aacattgaca gcaatgc当地 atgtcccaac	10140
acaaaacaatg aatattttgc atcaaacaat gcacacatac aatacttacc ttacacccct	10200
acactcttcc agagtgttcc atcaagcagg aatctgtgaa gctggagaag cgggtgatct	10260
ttgaagggtt ggagtccaaa tgctactctg ttgagcccgt gctgcagtgc ctgccccggct	10320
gtatcccagt gagaaccact accgtcaacg ttggcttca ctgcctgccc agtgtgag	10380
tgcaagaaag ccgagcttct ttggaaaaca ttatctctgt tttggattt cgtgcagtct	10440
ccaatcacgt ctccttttc actgttgcag acacaactgt ggaccgttct ggtctgagca	10500
gcttccttga gaagagcatc gacctgaggg atactgcaga agcccacctg gcctgtcgct	10560
gcactcctca gtgtgcttaa tctcattgcc ttccgtcaca tgacattttg ttttctgttt	10620
taggcagggt tggttaatta gcttggctc aacctgtaaa cttaataaa ttggaaaagc	10680
atgtgaagca taaaaactgtt gtgctcatct ttcatatata ttccattattc ttattataaa	10740
taaaacataa ataagatcaa cttaattttgt cccaaaagtt agttacagaa accaaaattt	10800
aaaaaaaaatg caacatgaca aacgtcagta aggtcgagtc tactcaaaag aagcaaaaact	10860
aaatagaaaatg attaaagaaa agtaaaataa ataaaataaa tctaaatcaa atgcaataaa	10920
atgttatata caaagcttga tataatacag taaggagctc aaaatattaa aatcttaagg	10980
aaatgcattt ataaaaaaac aataaaaatg taaaataaaag tgatttaca gagtaagatt	11040
aaacttagata aaaatggatt actaaaacaa atctgctcat cctcttcttg tatgtacttg	11100
ttgcacattt tgcatgaca tggaaatctt gtaaacacag tagtttataa aattgcatt	11160
tgtatataaa ccttaagcagt ttaggaacat ttatcacggt acaagtggaa gagaccctt	11220
tactgcactc cttgattttgt agaagggttgg cttattcgaa tccttataat gtaatcaca	11280
cgggtcacat ctgtgggac aggatgctgc ggctcgctc atcctgcaat ctggaaatgt	11340
tcaccttcc ttttgcagaa ctgcacattt acattggct cttacatgca caccagtaag	11400
aaaggcgtgc tttagcttt aattcagagg gtttaacaaa aatcttgcatt cag	11453

//